Attorney's Docket No.:10559/195001/P8367/Intel Corporation

In the specification:

Please amend the paragraph beginning at page 3, line 17 as follows:

FIG. 1 is a flow diagram of a gesture recognition process 100. Process 100 includes segmenting video data into video clips based on timing data (block 102). The timing data is used to define a window within which a user is expected to perform a single desired gesture. The video data is segmented so that each video clip contains the video data for a single window. In one embodiment, the timing data is a function of an audio signal having a beat. As used here, beat refers to any audibly perceptible semi periodic pulse contained within an audio signal. In such an embodiment, the user is expected to perform various predefined gestures "on the beat." That is, the window in which the user is expected to perform each gesture is defined by the beats of the audio signal. For example, the window can be defined so as to require the user to perform a desired predefined gesture within one second after a beat is played by a speaker. Alternatively, the window can be defined so as to require the user to perform the desired gesture in a one second time period starting one-half second before a beat is played by the speaker and ending one-half second after the beat. Another alternative is to define the window by a pair of adjacent beats are of the audio signal.

3'

Please amend the paragraph beginning at page 5, line 16 as

follows:

FIG. 3 is a flow diagram of a process 300 of determining

the probability that a video clip contains a predefined gesture.

First, movement of the user's body is identified and tracked for

each frame in the video clip (block 302). In one embodiment,

the moving regions in each video frame in the video clip are

identified. A three frame difference classifier can be used to

identify the moving regions in each video frame in the video

clip. For a given video frame in the video clip (referred to

here as the current frame), a pixel-by-pixel comparison of the

current frame and the immediately preceding frame, and a pixel-

by-pixel comparison of the current frame and the immediately

following frame, is performed. For a given pixel in the current

frame, the difference between the color of that pixel and the

color of the corresponding pixel in the immediately preceding

frame is determined. Also, for a given pixel in the current

frame, the difference between the color of that pixel and the

color of the corresponding pixel in the immediately following

frame is determined. If both differences exceed a predefined

tolerance for a given pixel, that pixel is considered to be a

moving pixel. The moving pixels in the current frame are then

clustered into moving regions (also referred to as "blobs")

4

using morphological image processing operations.   The

morphological image processing operations operate to remove

noise from the blobs.   Alternatively, other motion estimation

and clustering techniques can be used to identify the moving

regions in each frame of the video clip.

5